Recall: The definition of covariance is $\text{Cov}(X, Y) = \text{E}[(X - \mu_X)(Y - \mu_Y)]$.

Recall: The variance of a sum is $\text{V}(aX + bY) = a^2\text{V}(X) + b^2\text{V}(Y) + 2ab\text{Cov}(X, Y)$.

Recall: The sample mean is $\bar{Y} = \sum_{i=1}^{n} \frac{Y_i}{n}$.

Recall: The sample variance is $S^2 = \frac{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}{n-1} = \frac{\sum_{i=1}^{n} Y_i^2 - n\bar{Y}^2}{n-1}$.

Recall: The exponential pdf is $f(t) = \lambda e^{-\lambda t}$. The corresponding cdf is $F(t) = 1 - e^{-\lambda t}$ for $0 \leq t$. The corresponding mean and standard deviation are $1/\lambda$.

Recall: The second two homework assignments were based on $T = X_1 + \max\{X_2, X_3\}$, where $X_i$ is exponential with means $\mu_1$, $\mu_2$, and $\mu_3$, respectively. In HW2, $X_1$, $X_2$, and $X_3$ were independent, with standard errors estimated with micro-macro replications. In HW3, the effect of correlation between $X_2$ and $X_3$ was considered.
1. For truly random numbers, the expected value is 0.5 and variance is $1/12$ and lag-1 autocorrelation is zero. In a Monte Carlo evaluation (HW2) of the 16807 random-number generator, a student reported

- Expected Value = 0.500313
- Standard error of Expected Value = 0.000001
- Variance = 0.083252
- Standard error of Variance = 0.0000003
- Standard Deviation = 0.288535
- Lag-1 Correlation = -0.000012
- Standard error of Lag-1 Correlation = 0.000250

(a) (4 points) Circle the values that would change values if the experiment were rerun with a different random number seed. << circle all >>

(b) (6 points) Discuss the implications of the first two lines of the student’s results.

The two lines are "Expected Value = 0.500313" and "Standard error of Expected Value = 0.000001". The second line implies that the estimated standard error is significant through all digits of 0.500313. If so, the experiment is showing a statistically significant deviation from the mean of truly random numbers, which is 0.5. Therefore, there is a bug in the implementation of U16807 or in the implementation of the experiment code.

(c) (6 points) Assuming that the experiment was correctly coded and used $n$ random numbers, determine the value of $n$.

The standard deviation of truly random numbers is $\sqrt{1/12}$, which is about 0.29, as shown in the fifth line. The standard error of the expected-value estimator is 0.000001, as shown in the second line. Assuming independence implies that $\text{ste}(\bar{X}) = \text{std}(X) / \sqrt{n}$. Substituting the values yields

$$n \approx \left(\frac{\text{std}(X)}{\text{ste}(\bar{X})}\right)^2 = [0.28 / 0.000001]^2 = (290,000)^2 \leftarrow$$

Comment: Since the implied sample size is larger than $2^{31} - 1$, we have another indication that there is an error in the code.

(d) (6 points) There is no standard error reported for "Standard Deviation". I wonder "why not?". Circle the reason(s) that you can defend. (For partial credit for wrong answers, on the back of this page write a short defense for each of your choices.)

<< None are defendable >>

(i) The standard error of the standard deviation is obvious: it is the same as the standard error of the variance.

(ii) The standard error of the standard deviation is obvious: it is zero.

(iii) The standard error of the standard deviation has no micro-macro estimator.
2. For a single linear control $C$, the control-variate estimator is $\hat{\theta}(\alpha) = \hat{\theta} - \alpha(C - \mu_C)$.

The optimal value of $\alpha$ is $\alpha^* = \frac{\text{Cov}(\hat{\theta}, C)}{\text{Var}(C)}$.

Consider a Monte Carlo experiment that estimates $E(T)$, where $T = X_1 + \max\{X_2, X_3\}$ and $X_1$, $X_2$, and $X_3$ are independent and exponential with means $\mu_i$ for $i = 1, 2, 3$. Assume that the experiment uses $k$ independent macro-replications of $m$ micro-replications.

(a) (6 points) Suggest a control $C$ for this application.

Various control variates are reasonable. The mean must be known and the control variate correlated with $T$, the sample average of the random variable $T$.

For example, $\bar{X}_1$, the sample mean of $X_1$.
For example, $\bar{X}_2$, particularly if $E(X_2) > E(X_3)$.
For example, $\bar{X}_2 + \bar{X}_3$.
For example, $\bar{X}_1 + \bar{X}_2 + \bar{X}_3$.

Comment: Answers to Parts (b), (c), and (d) depend upon the answer to Part (a).

(b) (4 points) (circle one of the four choices) For your choice of $C$, will $\alpha^*$ be

(i) negative (ii) zero (iii) positive ←
(iv) so random that the sign of $\alpha^*$ is only known after the experiment is run.

(c) (5 points) In estimating $\alpha^*$, $\text{Var}(C)$ needs to be estimated. Suggest a point estimator of $\text{Var}(C)$.

The observations are independent, so the grand estimator can be used.

$$S_C^2 = \frac{\sum_{i=1}^m \sum_{j=1}^k C_{ij}^2 - n \bar{C}^2}{n (n - 1)}$$

where $n = km$.

Alternatively, the micro-macro estimator can be used.

$$S_C^2 = \frac{\sum_{j=1}^k S_{Cj}^2}{k}$$

where $S_{Cj}^2 = \frac{\sum_{i=1}^m C_{ij}^2 - m \bar{C}_j^2}{m - 1}$

(d) (3 points) T F ← To induce a non-zero correlation between $\hat{\theta}$ and $C$, common random numbers are needed.
3. (3 points each) Based on the notation of Problem 2, circle one of "constant", "random variable", "event" or "meaningless".

(a) \( \mu_C \)  
constant \( \leftarrow \) random variable \( \rightarrow \) event \( \rightarrow \) meaningless

(b) \( \alpha^* \)  
constant \( \leftarrow \) random variable \( \rightarrow \) event \( \rightarrow \) meaningless

(c) \( \text{Cov}(\hat{\theta}, C) \)  
constant \( \leftarrow \) random variable \( \rightarrow \) event \( \rightarrow \) meaningless

4. (3 points each) The acceptance-rejection method accepts a candidate value \( x \) from \( r(x) \) with probability \( 0 \leq f(x) / t(x) \leq 1 \).

(a) T \( \leftarrow \) F The quantity \( x \) can be a vector.

(b) T \( \leftarrow \) F The values of \( x \) can be discrete or continuous.

(c) T \( \leftarrow \) F If the values of \( x \) are continuous, then the probability that the candidate value is accepted is \( 1 / \int_{-\infty}^{\infty} t(x) \, dx \).

5. (3 points each) The composition method considers \( f(x) = \sum_{i=1}^{k} p_i f_i(x) \). A random variate is generated from \( f_i \) with probability \( p_i \).

(a) T \( \leftarrow \) F The quantity \( x \) can be a vector.

(b) T \( \leftarrow \) F The values of \( x \) can be discrete or continuous.

(c) T \( \leftarrow \) F Generation from \( f_i \) must be via the inverse cdf transformation.
6. Consider the Poisson process with rate function $\lambda(t) = 6t^2$. Then the cumulative rate is $\Lambda(t) = 2t^3$. At time $t_{i-1}$, the cdf of the time of the next arrival is $F(t_i) = 1 - \exp(-((\Lambda(t_i) - \Lambda(t_{i-1}))))$.

(a) (6 points) Suppose that time is now $t = 4.5$ and that the next random number is $u = 0.4$. Determine the time of the next arrival.

Two solutions. Let $T_i$ denote the time of the next arrival. Then $t_{i-1} = 4.5$.

1. Compute the inverse cdf:
   The conditional cdf is
   
   
   $F_{T_i}(t_i) = 1 - \exp[-(2t_i^3 - 2(4.5)^3)]$
   
   for $t_i \geq 4.5$ (and zero elsewhere).

   Setting $u = 0.4$ to the cdf and solving for $t_i$ yields
   
   $t_i = [(4.5)^3 - 0.5 \ln(1-0.4)]^{1/3}$.

2. Compute exponential area.

   Compute the standard exponential random variate $a = -\ln(1 - u) = \ln(0.6)$. Compute the time of the next arrival as the time that provides area $a$ under $\lambda$ to the right of $t = 4.5$.

(b) (6 points) For this application, do you prefer thinning or the inverse cdf transformation? Explain why.

Because the rate goes to infinity, thinning is difficult. There is no constant rate $\lambda_*$. A non-constant rate would need to be similar to $\lambda$ so using the inverse cdf transformation seems to be the better choice.
7. NORTA (Normal to Anything) is a method for obtaining $k$-dimensional random vectors in Monte Carlo simulation experiment. In Step 1, independent standard normal random variates are converted to correlated standard normal random variates, usually via the Cholesky decomposition. In Step 2, the correlated standard normal random variates are converted to correlated $U(0,1)$ random variates, via the standard normal cdf. In Step 3, random variates with the specified properties are obtained from the correlated $U(0,1)$ random variates, via the inverse cdf transformation.

(a) (3 points) T ← F ← In HW3, the values of corr($X_2, X_3$) range from −1 to 1.

(b) (3 points) T ← F ← In HW2, the value of corr($X_2, X_3$) is maximized.

(c) (3 points) T ← F ← In HW3, the value of $E(T)$ is a constant, and therefore does not change value as a function of corr($X_2, X_3$).

(d) (3 points) T ← F ← In NORTA, the correlations of Step 1 have the same values as the correlations of Step 3.

(e) (3 points) T ← F ← In NORTA’s Step 3, acceptance-rejection could be used (rather than the inverse cdf transformation) if the random numbers are synchronized using multiple random number streams in conjunction with common random numbers.

(f) (6 points) Discuss how your C code computed the standard normal cdf in Step 2.

There are many approximations to the standard normal cdf. Internet search for an appropriate C code is easy.

This question is designed to check whether the NORTA homework (HW3) was done.